

Enhanced target recognition employing spatial correlation filters and affine scale invariant feature transform

Akber Gardezi^{*a}, Usman Malik^b, Saad Rehman^c, R.C.D. Young^d, P.M. Birch^d, C.R. Chatwin^d

^a COMSATS University Islamabad, Pakistan; ^b Center for Advanced Studies in Engineering, Islamabad, Pakistan, ^c EME College, National University of Science and Technology, Islamabad, Pakistan, ^d Department of Engineering and Design, University of Sussex, Brighton, United Kingdom

ABSTRACT

A spatial domain optimal trade-off Maximum Average Correlation Height (SPOT-MACH) filter is an advanced and newly developed method which has shown to have advantages over frequency domain implementations of the Optimal Trade-Off Maximum Average Correlation Height (OR-MACH) filter as it can be made locally adaptive to spatial variations in the input image background clutter and normalized for local intensity changes. This enables the spatial domain implementation to be resistant to illumination changes. The Affine Scale Invariant Feature Transform (ASIFT) is an extension of previous feature transform algorithm and is a fully affine invariant algorithm; its features are invariant to six affine parameters which are translation (2 parameters), zoom, rotation and two camera axis orientations. This results in attaining strong features thus giving accurate and increased key points which can then be used for matching between different images of the object being tested. In this paper a novel approach will be adopted for enhancing the performance of spatial correlation filter (SPOT MACH filter) using ASIFT in a pre-processing stage enabling fully invariant object detection and recognition in images with geometric distortion/ alteration. An optimization criterion would also be developed to overcome the temporal overhead of the spatial domain approach.

In order to evaluate effectiveness of algorithm, experiments were conducted on two different data sets. Several test cases were created based on illumination, rotational and scale changes in the target objects.

The performance of correlation algorithms was also tested against composite images as reference and it was found that this results in a well-trained filter with better detection ability even when the target object has gone to large rotational changes.

Keywords: MACH filter, SIFT filter, correlation filter, invariant pattern recognition

1. Introduction

Previously Gardezi et.al identified the major drawbacks of frequency domain correlation filters [5] which were prone to giving false detections in the presence of non-uniform lighting conditions. However, the requirements for an optimal correlation filter is the invariance to multiple affine transformation parameters such as zoom, rotation, and translation. This is considered a prerequisite for all object recognition systems based on real-time data acquisition and processing for security, and surveillance systems.

The Scale Invariant Feature Transform (SIFT) algorithm [3] is based upon a concept of feature detection and it provides a feature-based description of the target in an image by extracting a group of interest points. The extracted feature description from the image can be used for detecting the target in a cluttered/ multiple object scene. The matches between the target and reference (input) image are identified by finding the two nearest neighbors of each key point from the first image among those in the second image, and only considering a match if the distance to the closest neighbor is less than the

* akber.gardezi@comsats.edu.pk; phone +92 51 8731978; fax +92 9247006

threshold value of the second closest neighbor. The threshold can be increased to increase matches but there is a compromise on reliability.

The neighboring approach enables the SIFT algorithm to be tolerant with changes in scaling, orientation and partially invariant to distortion and illumination changes which can be used for identifying objects from cluttered backgrounds [2]. However, the invariance to out-of-plane rotation of SIFT is severely compromised in case the 90-degree rotation over the Y- Axis which causes the front of the object to be pressed across the X- Axis.

One of the solutions to tackle such an out-of-plane rotation is ASIFT. The ASIFT simulates all the views of the image that can be obtained by changing the latitude and longitude angles. Afterwards it employs the basic SIFT method for completion of the process. This additional capability enhancements results in ASIFT to be invariant to out-of-plane rotations like Figure 1 given below:



Figure 1. (a) Image with no Rotation, (b) Image with Y-Axis out of plane rotation

In this paper a system is presented that employs ASIFT [2] followed by the Spatial domain Optimal Trade off Maximum Average Correlation Height (SPOT-MACH) filter [5] for tolerance to illumination changes.

The desired result in this case is a fully invariant object detection and recognition system. However, the time overhead owing to computational intensity has been addressed through optimization techniques to enhance the speed of the system.

2. Hybrid Design Approach

The hybrid image recognition technique proposed in this paper combines the efficiency and accuracy of two algorithms (SPOT –MACH and ASIFT) in order to achieve better recognition in varied lighting conditions.

The adopted approach uses ASIFT as a preprocessor to match the key points between the reference and target images and in case the match exceeds the threshold a positive object recognition is achieved. Once a match is received further processing is performed to save the computational overhead.

The process flowchart of the proposed method is shown in figure 2 given below:

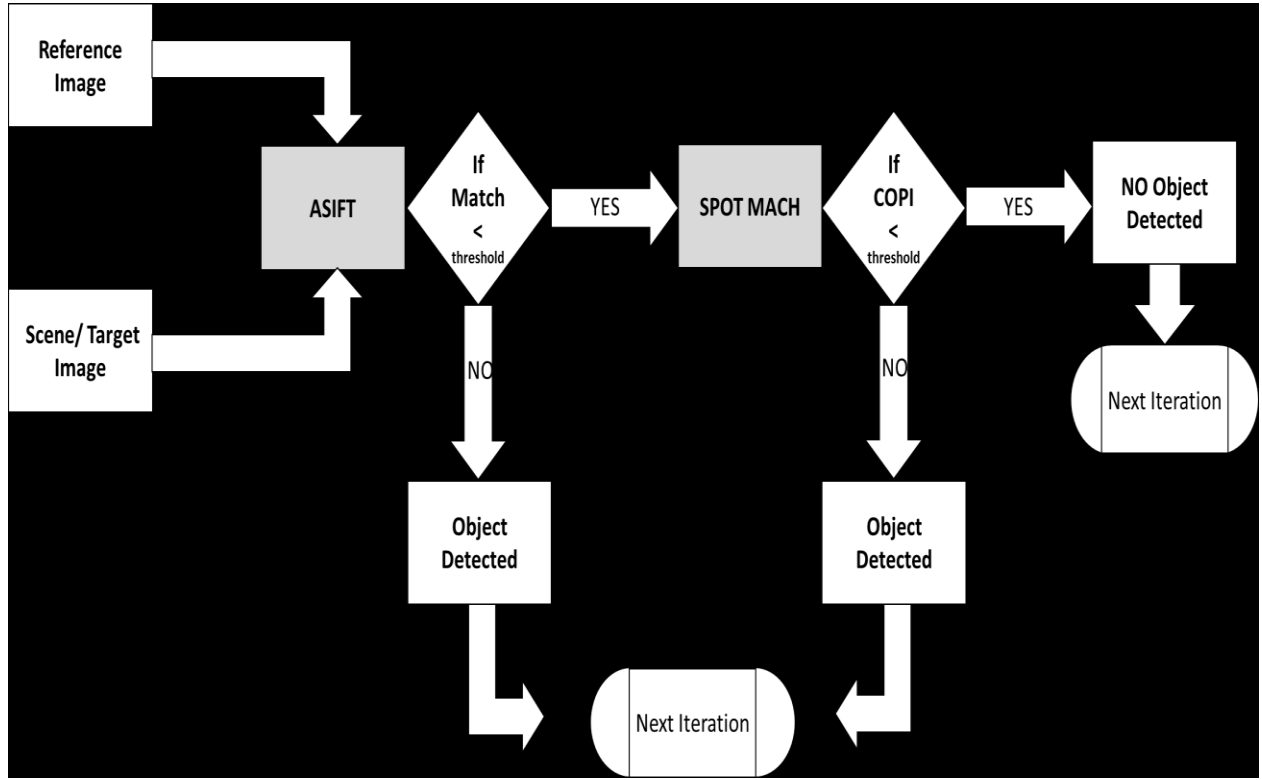


Figure 2. Process flowchart of hybrid image recognition system

However, in the case where the ASIFT fails to generate matches less than or equal to the threshold than the SPOT-MACH filter is applied. In this case the resultant Correlation Output Peak Intensity (COPI) value will determine the result of detection.

It may be noted that target image in this case is proposed to be a cropped image based on the matchings of ASIFT algorithm. General area around the most matchings cropped from the scene to be passed on to the SPOT MACH filter. This will not only save computational overheads but also result in improved accuracy of the filter.

The SPOT-MACH filter enables adaptation of the filter dependant on locally variant background clutter conditions and also enables the normalisation of the filter energy levels at each step.

Thus, the SPOT-MACH filter offers an advantage over its frequency domain implementation as shift invariance is not imposed upon it. The only drawback of the spatial domain implementation is the amount of computational resources that are required for its real time implementation.

In order to formulate the transfer function for the SPOT-MACH filter, first the OT-MACH transfer function should be computed in the frequency domain, after which the reference image spectrum is inverse Fourier transformed to yield a space domain image as shown in Equation (1):

$$h^* = IFFT \left[\frac{m_x^*}{\sqrt{\alpha C + \beta D_x + \gamma S_x}} \right] \quad (1)$$

Where the terms α , β and γ , are the non-negative weighting parameters; m_x which is the average of the training image vector x_1, x_2, \dots, x_N (in the frequency domain); C which is the diagonal power spectral density matrix of additive input noise and is usually set as the white noise covariance matrix, $C = \sigma^2 I$; D_x is the diagonal average power spectral density of the training images [1].

Furthermore, in the case of SPOT-MACH filter the size of the correlation output plane is smaller as compared to the frequency domain plane. This is due to the windowing function used resulting in returning a lesser number of correlation points if zero padding is not used. The correct size of the correlation plane can be established from equation 2:

$$Corr_plane_size(x, y) = (x - window_size, y - window_size) \quad (2)$$

Since the correlation output plane size has been compensated, the location of the peak can be made to correspond to the location of the object in the input scene. Thus, in order to determine the corrected location of the target object, the following transformation should be applied to the peak coordinates of the SPOT-MACH filter:

$$Corr_Outputplane = \left(x - \frac{winsize}{2}, y - \frac{winsize}{2} \right) \quad (3)$$

The use of Equation (3) will determine the exact location of the object in the correlation output plane for the SPOT-MACH filter. The use of a moving kernel enables the filter to be adaptive to small areas of the image and hence minimise the impact of varying illumination patterns. Once the images are passed through the SPOT-transfer function they are correlated using a spatial correlator. In this approach a windowing kernel normally the size of the target object is extracted from the composite image created by the SPOT-MACH and is scanned through the target image pixel by pixel. This enables adaptation of the filter dependant on background illumination variations and also enables the normalisation of the filter energy levels. The kernel can be normalised to remove a non-uniform brightness distribution if this occurs in different regions of the image. The main constraint in this implementation is the dependence on computational ability of the system.

In this paper the Amsterdam Library of Object Images (ALOI) was selected for experimentation owing to high quality RGB imagery data set of up to 110,250 images from a collection of 1000 different objects. The images were chosen to capture the sensory variation in object recordings based on varied viewing angle (rotation), illumination angle, and illumination color for each object. The primary interest in this case was in viewing angle and the illumination angle. From the object database two images were selected for testing and experimentation as shown in figure 3. The image having the 'small shoe' was chosen due to its irregular shape and the other image having the cylindrical box of chips was chosen due to its regular geometric shape.



Figure 3. (a) Irregular Shape Reference Image - Shoe, (b) Regular Shape Reference Image – box

3. Discussion and Results

In this section the different scenarios were created to test the capabilities of the OT-MACH filter in the frequency domain, SPOT-MACH, SIFT and ASIFT filters. The aim was to identify the tolerance of these filters when exposed to non-uniform lighting conditions. A set of objects were chosen from the Amsterdam image laboratory and tests were conducted to identify the target object.

Initially to assess the illumination tolerance factor of the frequency domain and spatial domain approaches a set of composite images were created comprising of a regular shaped object and irregular shaped object. The composite images used in the simulations composed of the reference images which were rotated at 0 to 90 degrees with a step size of 10 degrees each for an irregular shaped object are shown below in figure 4.

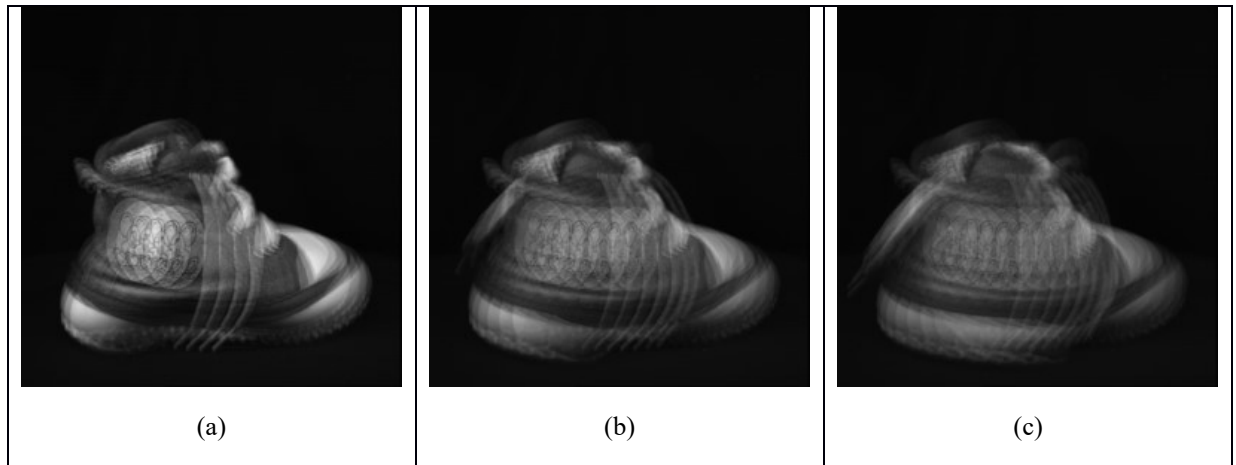


Figure 4. Composite images for irregular shape (a) 0 to 30 degrees, (b) 0 to 60 degrees, (c) 0 to 90 degrees

The composite images used in the simulations composed of the reference images which were rotated at 0 to 90 degrees with a step size of 10 degrees each for a regular shaped object are shown below in figure 5.

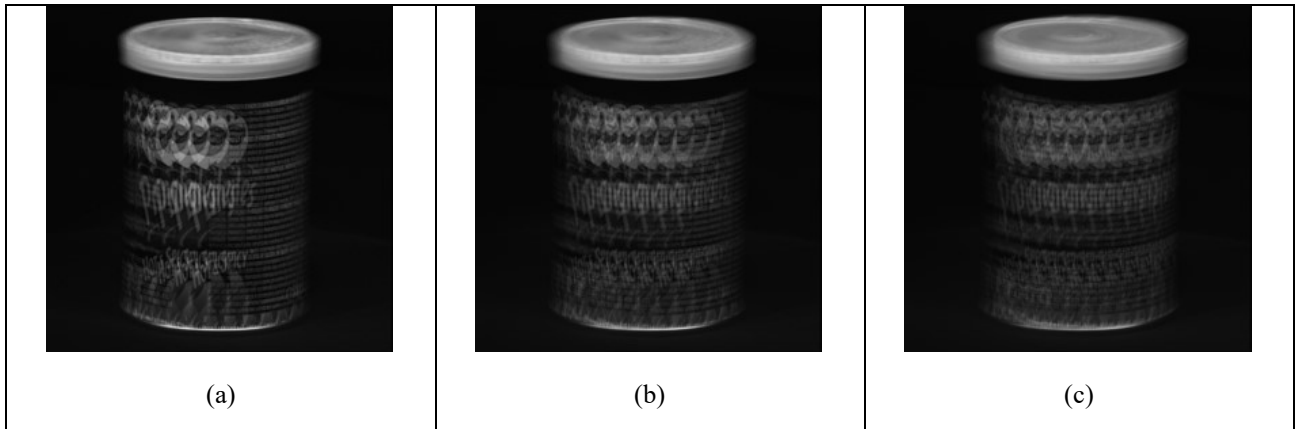


Figure 5. Composite images for regular shape (a) 0 to 30 degrees, (b) 0 to 60 degrees, (c) 0 to 90 degrees

A set of varying illumination shades were created to test the tolerance in illumination changes for the OT-MACH, SPOT-MACH and ASIFT implementations. The first test image is given by Figure 6 below:

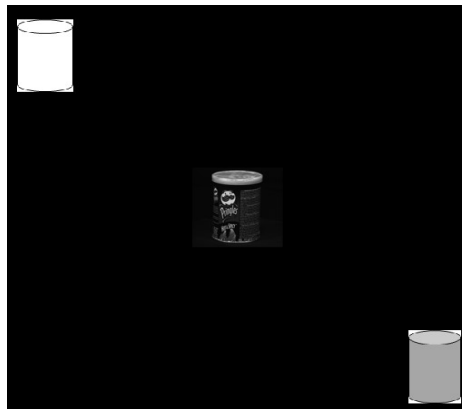


Figure 6. Target image for varying illumination gradient for a regular shaped object

The main aim of this test is to assess the tolerance of the SPOT-MACH filter to the change in illumination as compared to MACH Filter when one or more false targets with varying illumination gradients are present.

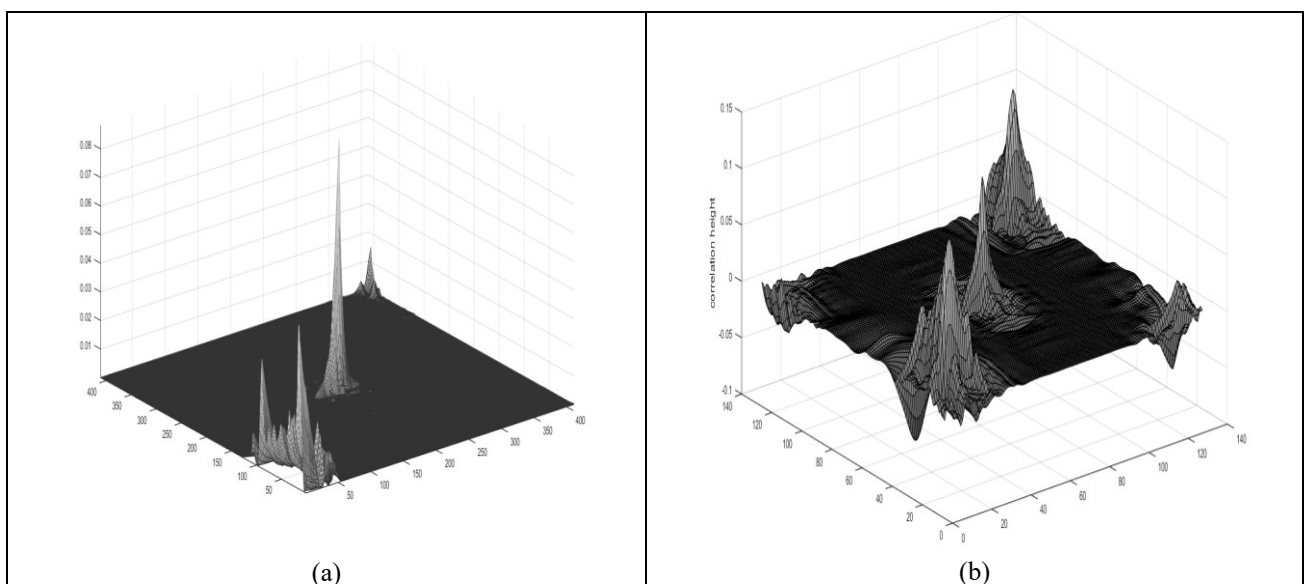


Figure 6. (a) Correlation output: SPOT-MACH, (b) Correlation output: OT-MACH

From the figure 6 given above it is quite evident that only SPOT MACH was able to cater for the false targets which were presenting a similar shape and bright illumination characteristics.

The OT-MACH filter's inability to distinguish between such false targets is clear from figure 6 (b) as the correlation plane shows three peaks of similar height whereas the SPOT MACH, in figure 5 (a) manages to identify the real target at its correct location and gives a sharp correlation peak.

Similarly, the ASIFT also gives a confirmed result with 257 matches and no false given by figure 7 below:

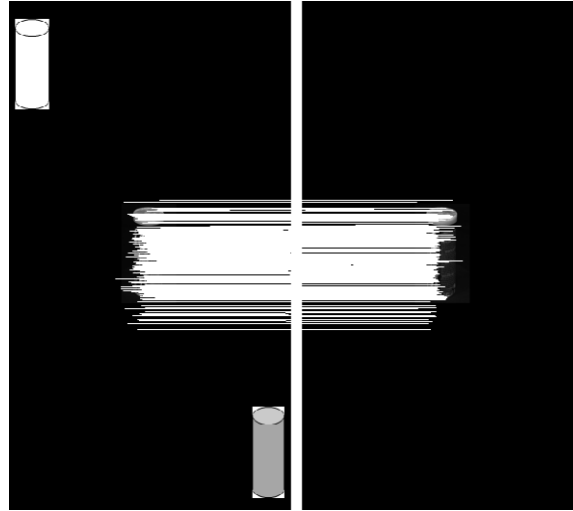


Figure 7. ASIFT Matches for varying illumination gradients

The above experiments show the illumination change tolerance of ASIFT, OT-MACH and SPOT-MACH filter. The parameters used for alpha, beta and gamma for OT-MACH in the experiments are 0.001, 0.35 and 0.1 respectively. In the case of parameters for the SPOT-MACH filter, the values for alpha, beta and gamma used are 0.01, 0.7 and 0.01 respectively.

3.1 Experiment 1 - Using Irregular Shaped Reference Object

In this case multiple algorithms, SIFT, ASIFT, OT-MACH and SPOT MACH were tested against the irregular shape object (shoe) chosen for this experiment. The lighting conditions in the data set were normal and only the images with rotational variations were used. The target image used in this experiment is shown in figure 8 below:

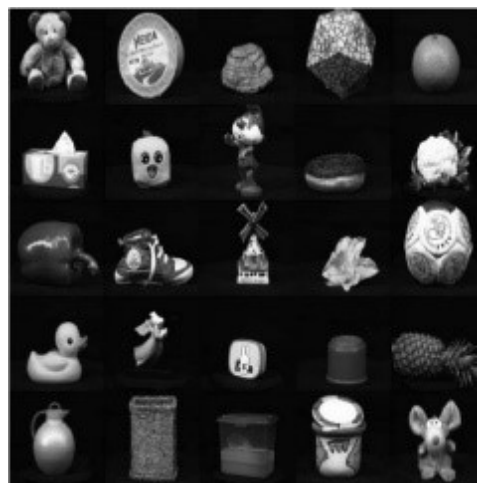


Figure 8. Target image for irregular shaped object experiment

In the first instance, the SIFT and ASIFT were tested using different angle variations shown by Figure 9 below:

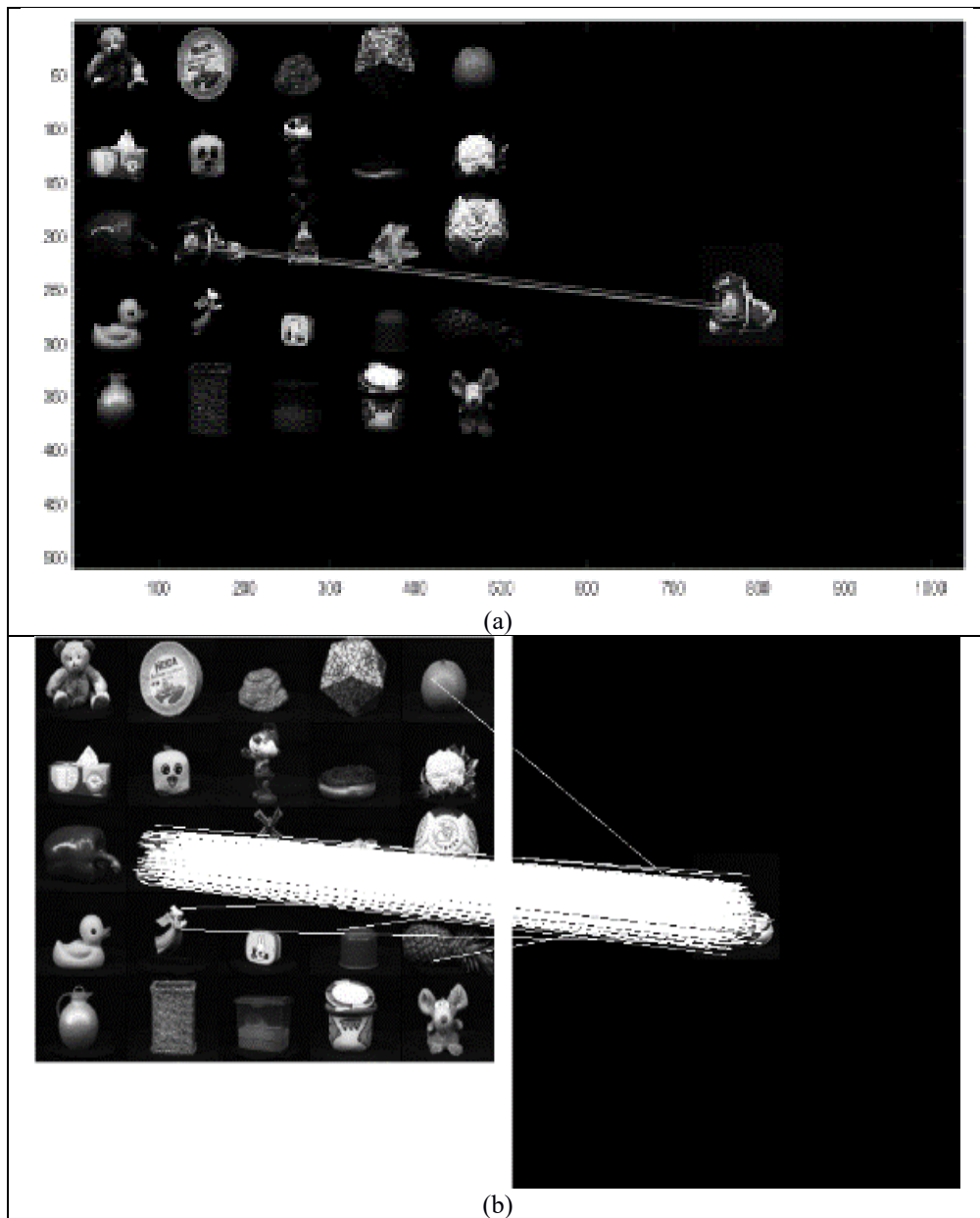


Figure 9. (a) SIFT Result for irregular shaped object (shoe), (b) ASIFT Result for irregular shaped object (shoe)

From Figure 9 it can be seen that the SIFT algorithm makes lesser matches than the ASIFT. However, although the number of matches in ASIFT are more, but some false matches are also present as can be seen from Figure 9(b).

Furthermore, when the illumination profile is changed the number of false matches increases in the case of ASIFT which can be seen from Figure 10.

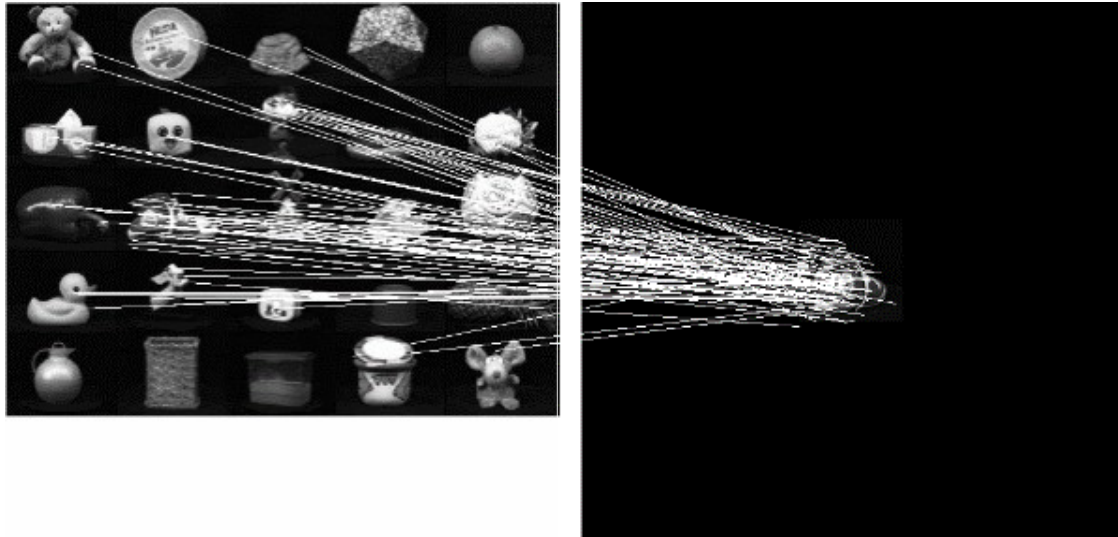


Figure 10. ASIFT results for irregular shaped object (shoe) in bright illumination profile

However, in order to overcome the problem of illumination changes the hybrid approach presented in Figure 2 was used where the region with the most matches was extracted and passed through SPOT-MACH to detect the object as shown by Figure 11 below:

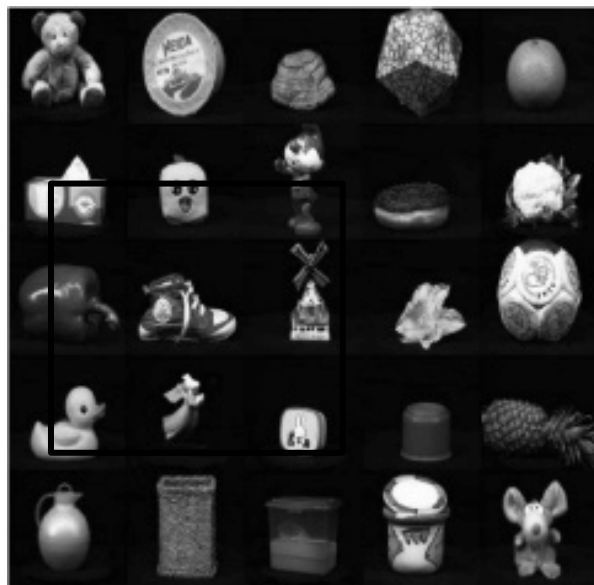


Figure 11. Extracting region of interest from ASIFT matches

The region of interest was extracted from the target image with the bright illumination profile and was subsequently filtered by OT-MACH and SPOT filter to assess the detection capabilities in frequency and spatial domain.

The results of the OT-MACH and SPOT-MACH filter the region extracted from the bright illumination profile are given by Figure 12 and 13 below:

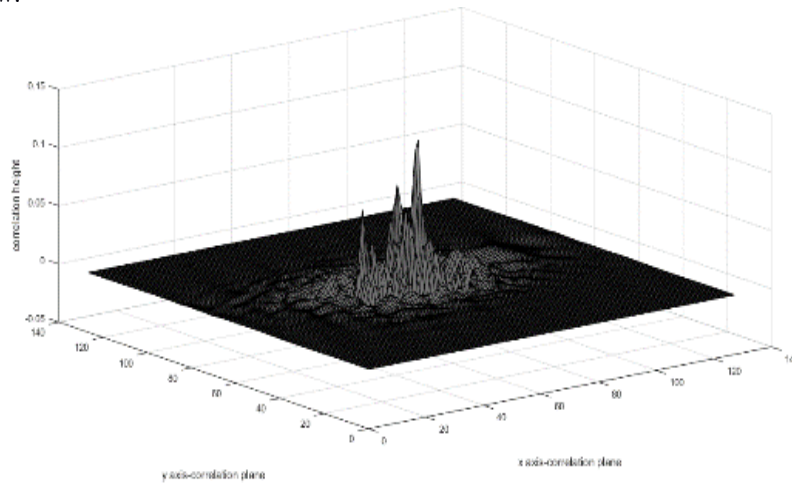


Figure 12. OT-MACH correlation plane for bright illumination

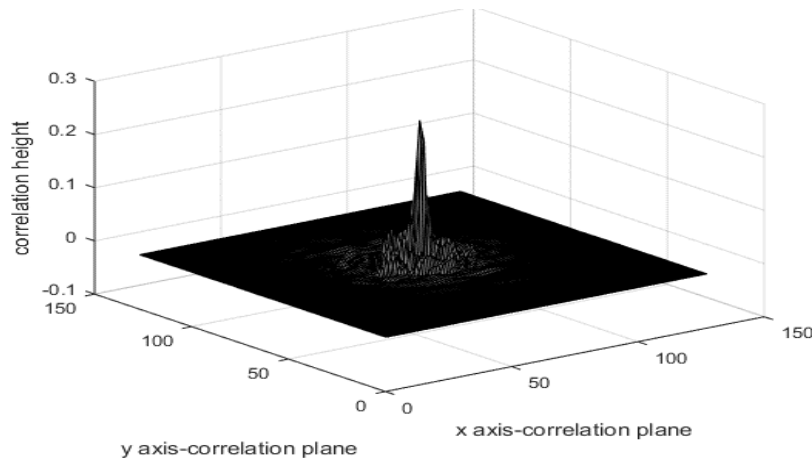


Figure 13. SPOT-MACH correlation plane for bright illumination

From Figure 12 and 13 it can be seen that the SPOT-MACH filter gives a better detection when a bright illumination gradient is present in the image.

From the experiment results obtained for an irregular shaped object to determine the capability of the selected algorithms, it can be deduced that:

- i. Performance of SIFT algorithm degrades in case of illumination and orientation changes;
- ii. ASIFT performs well and up to 30 degrees rotation of the target image with gradual degradation between 40-50 degrees;
- iii. The OT-MACH filter performance is good up to 30-degree orientation changes;
- iv. The SPOT-MACH performs better than OT-MACH beyond 50-degree orientation changes.

From the test results it is proposed that the ASIFT should be used as a preprocessing stage for determining the region of interest and then SPOT MACH will be used for completing the recognition phase.

A detailed outlay of the results generated during experiment is presented by Table 1 given below:

Sr.	Rotation of Target	SIFT (No. of matches)	ASIFT (No. of matches)	MACH Normalised COPI Value	SPOT MACH Normalised COPI Value
1	0	5	257	0.88	0.9894
2	10	1	147	0.77	0.741
3	20	4	147	0.7	0.714
4	30	1	87	0.64	0.652
5	40	0	80	0.55	0.697
6	50	0	76	0.52	0.616
7	60	0	71	0.51	0.065
8	70	0	62	0.45	0.0291
9	80	0	74	0.42	0.023
10	90	0	62	0.2	0.0257

Table 1. Outlay of results from experiment 1 using irregular shaped object for bright illumination

3.2 Experiment 2 - Using Regular Shaped Reference Object

Similar to experiment 1, the performance of all algorithms, SIFT, ASIFT, OT MACH and SPOT MACH was tested using a regular shaped object (box).

Initially the tests were conducted using uniform lighting conditions with rotational variations from 0 to 90 degree, and a step size 10 degree.

The target image used in this experiment can be seen in Figure 14 given below:



Figure 14. Target image for regular shaped object (box) experiment

Similar to experiment 1 the tests were conducted initially with SIFT and ASIFT algorithms in uniform lighting conditions given by Figure 15.

It can be seen from Figure 15, that the performance of ASIFT is better than SIFT from the perspective of false matches.

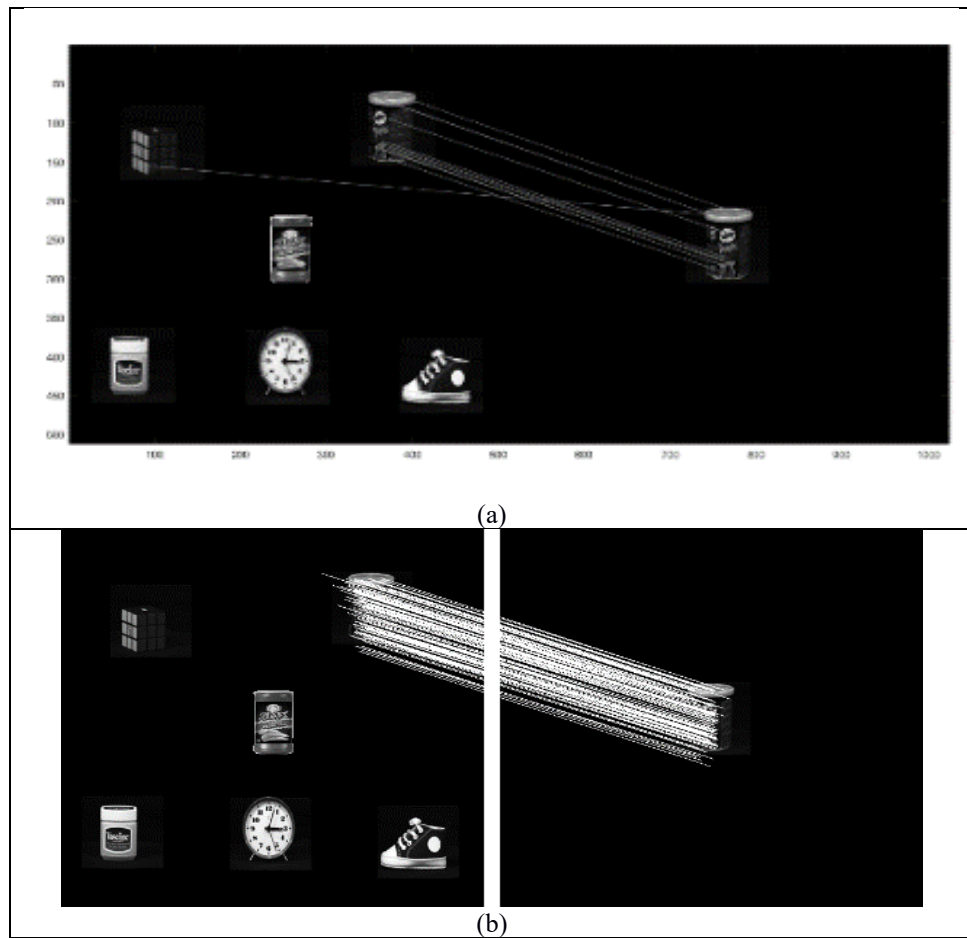


Figure 15 (a) SIFT Result for regular shaped object (box), (b) ASIFT Result for irregular shaped object (box)

However, when the bright illumination profile is used in the case of ASIFT the number of false matches increase as can be seen from Figure 16 below:

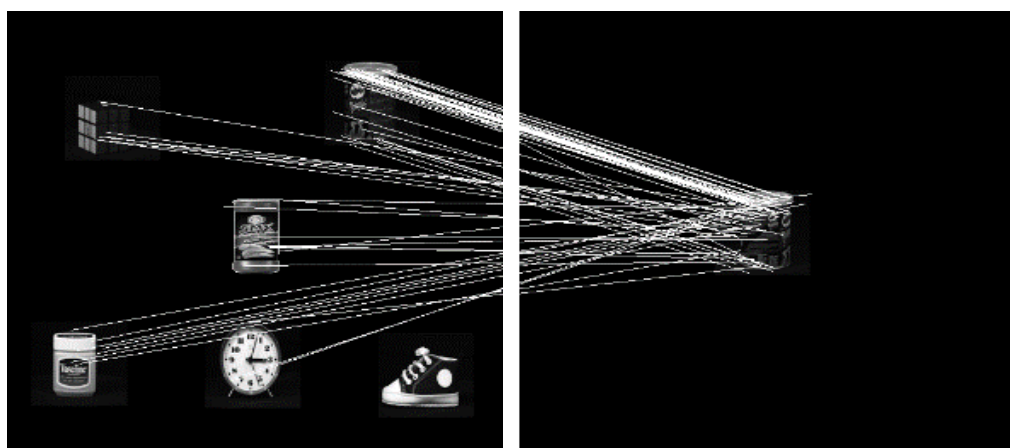


Figure 16. ASIFT results for regular shaped object (box) in bright illumination profile

Subsequently, the tests were conducted with SPOT-MACH and OT-MACH filter to assess the performance in the presence of bright illumination. However, in this case the cropped region of interest was not used as the false matches were spread over different parts of the target image. The correlation output plane for SPOT-MACH and OT-MACH filter are given by Figure 17 and 18 below:

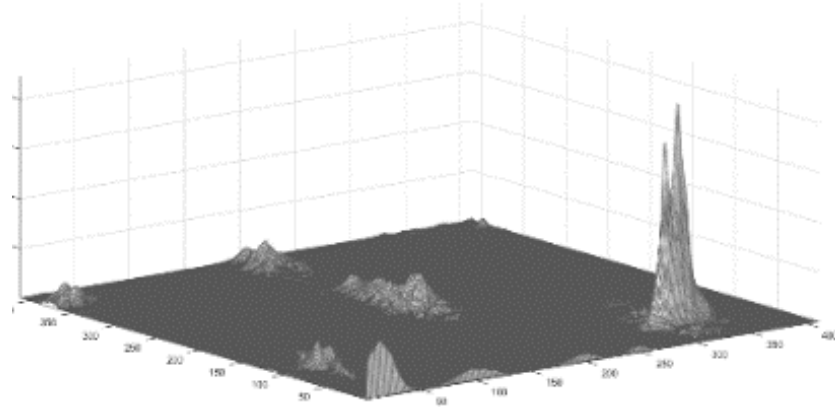


Figure 17. SPOT-MACH correlation plane for regular shape (box) object in bright illumination

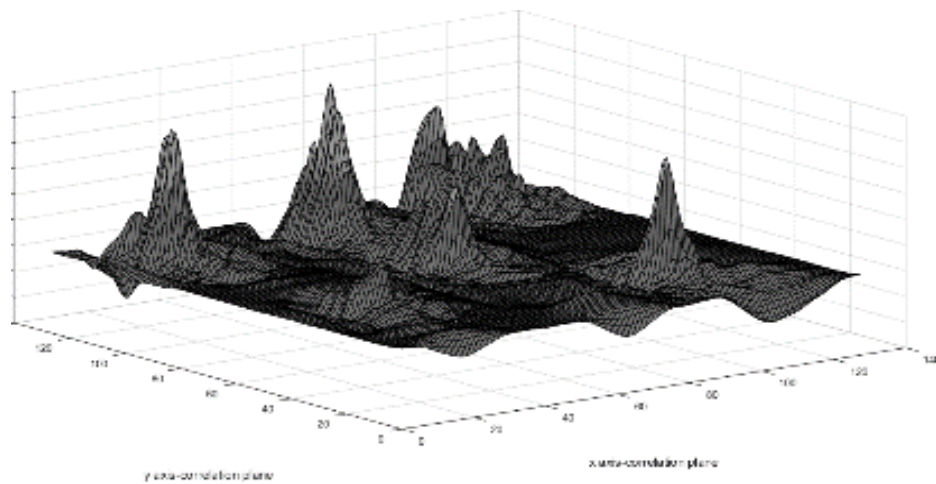


Figure 18. OT-MACH correlation plane for regular shape (box) object in bright illumination

From Figure 17 and 18 it can be seen that the SPOT-MACH filter is able to detect the location of the target object whereas the OT-MACH filter fails to give a detection in this case.

From the experiment results obtained for a regular shaped object to determine the capability of the selected algorithms, it can be deduced that:

- i. Performance of SIFT algorithm is better for regular shaped objects as compared to irregular shaped objects with increased number of matches and tolerance to orientation changes up to 30 degrees;
- ii. The ASIFT performance improved up to 50 degrees and performance degraded only after 70 degrees of rotation in the test image;
- iii. OT-MACH filter was not able to detect the real target in the scene;
- iv. The SPOT-MACH detected the target despite illumination variances as compared to OT-MACH filter;

An outlay of performance analysis of the chosen filters is given by Table 2 below:

Sr.	Rotation of Target	SIFT (No. of matches)	ASIFT (No. of matches)	MACH Normalised COPI Value	SPOT MACH Normalised COPI Value
1	0	40	295	0.78	0.9894
2	10	24	216	0.77	0.941
3	20	15	169	0.7	0.864
4	30	11	109	0.64	0.852
5	40	7	84	0.35	0.797
6	50	7	63	0.22	0.788
7	60	7	72	0.16	0.765
8	70	1	82	0.15	0.691
9	80	6	79	0.12	0.6
10	90	4	64	0.1	0.57

Table 2. Outlay of results from experiment 2 using regular shaped object for bright illumination

From experiment 1 and 2 the performance of SIFT and ASIFT is plotted given by Figure 19 below:

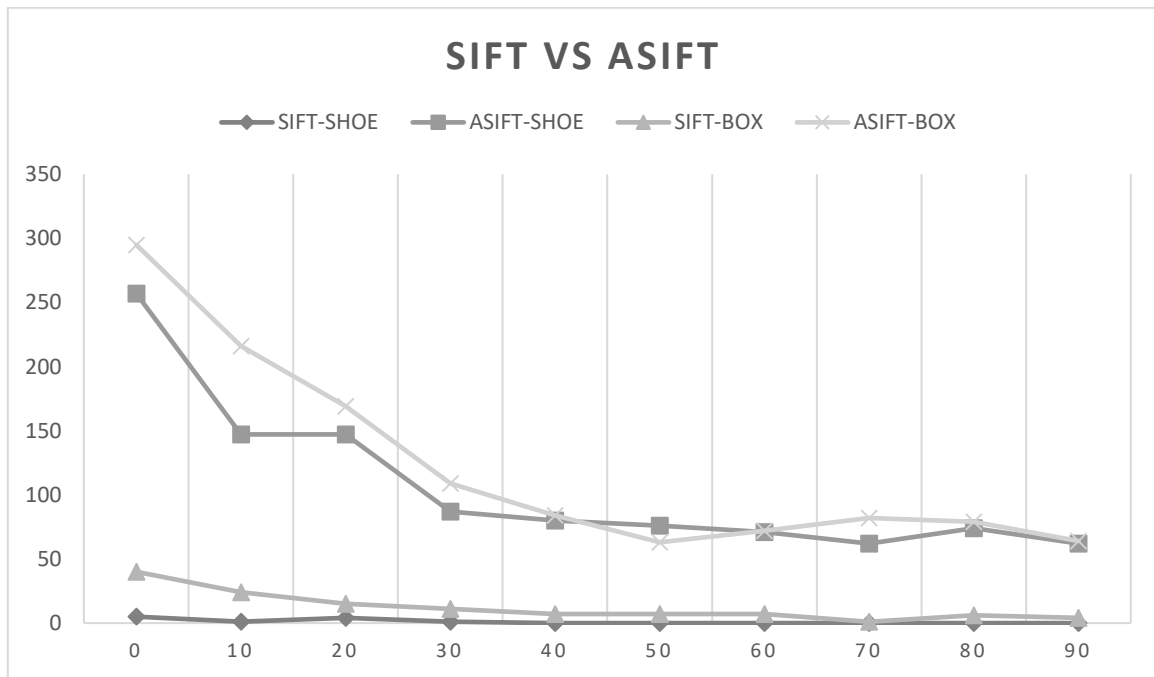


Figure 19. Performance chart of SIFT and ASIFT

From Figure 19 it can be seen that the performance of the SIFT and ASIFT degrades with the changes in orientations, however the ASIFT generates a greater number of matches over an extended range of orientation changes. Whereas in the case of SIFT the impact of orientation change on performance is evident from minor changes in orientation. The hybrid approach is best suited in this case for ASIFT as a preprocessor for better object recognition

4. Conclusion and Future Work

Owing to the computationally intensive nature of the SPOT-MACH THE Speed enhancement techniques need to be implemented thereby reducing the computational overhead. Further evaluation of the system for implementation of following applications is also recommended:

- i. Intrusion detection systems
- ii. Automatic number plate recognition
- iii. Prohibited item detection using X-ray images
- iv. Military grade video exploitation systems (post mission analysis)

In this paper correlation and non-correlation-based algorithms and techniques were implemented, tested and compared for their advantages and limitations. A detailed comparison was conducted of the SPOT-MACH filter with its frequency domain counterpart OT-MACH while SIFT and ASIFT algorithms were also tested for their effectiveness. A matrix was created for the comparison of each experiment; showing that the SPOT-MACH filter offers better discrimination ability when subjected to illumination variations while ASIFT algorithm can be used as a robust preprocessing algorithm for object recognition in most cases and filtering out region of interests in remaining cases. It was also discussed that the main advantages of the SPOT-MACH filter is the ability to allow localised normalisation of the filter which is not possible in the frequency domain implementation. The results presented which indicate that a spatial domain implementation of the filter is able to detect, locate and recognise a target object within a non-uniformly lit scene. The performance of the different methods was summarised in a table showing number of matches SIFT and ASIFT alongside COPI values for the correlation filters. It was observed that the performance of SIFT and ASIFT degrades when there are sudden changes in orientation and hence the SPOT-MACH filter can be used at the second stage for object recognition.

References

- [1] Mahalonobis, A. , Kumar, B.V.K. , Song, S., Sims, S.R.F. , Epperson, J.F., "Unconstrained correlation filters,". Applied Optics , 33 , 3751-3759(1994).
- [2] Morel, M. , Yu, G., "ASIFT: A new framework for fully affine invariant image comparison," . SIAM Journal on Imaging Sciences, 2(2), 438–469, (2009).
- [3] Lowe, D.G. " Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision 60(2), 91-110(2004).
- [4] Mahalonobis, A. , Kumar, B.V.K. , "Optimality of the maximum average correlation height filter for detection of target in noise," Optical Engineering 36(1), 2642-2648(1997).
- [5] Gardezi, A. , Birch , P. , Young , R. and Chatwin, C. , " Implementation of the Maximum Average Correlation Height (MACH) filter in the spatial domain for object recognition from clutter backgrounds," Proc. SPIE 7696, (2010).
- [6] Gardezi, A. , Al-Kandri, A. , Birch, P. , Young, R. and Chatwin, C. " A space variant maximum average correlation height (MACH) filter for object recognition in real time thermal images for security applications," Proc. SPIE 7838, (2010).
- [7] Gardezi, A. , Al-Kandri, A. , Birch, P. , Young, R. and Chatwin, C. , " Enhancement of the speed of space-variant correlation filter implementations by using low-pass pre-filtering for kernel placement and applications to real-time security monitoring," Proc. SPIE 8055, (2011).
- [8] Birch, P., Gardezi, A. ,Mitra,B.K., Young, R. and Chatwin, C., "An optical space domain volume holographic correlator," Proc. SPIE 7340,(2009).
- [9] Gonzalez,R.C, Woods, R.E., [Digital Image Processing], Prentice-Hall, (2002).